A dynamical systems perspective on driver behavior

Citation for published version:
Mirman, JH 2019, 'A dynamical systems perspective on driver behavior', Transportation Research Part F: Traffic Psychology and Behaviour, vol. 63, pp. 193-203. https://doi.org/10.1016/j.trf.2019.04.010

Digital Object Identifier (DOI):
10.1016/j.trf.2019.04.010

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Transportation Research Part F: Traffic Psychology and Behaviour

Transportation Research Part F 63 (2019) 193–203

Contents lists available at ScienceDirect

Transportation Research Part F

journal homepage: www.elsevier.com/locate/trf

A dynamical systems perspective on driver behavior

Jessica Hafetz Mirman

The University of Alabama at Birmingham, Department of Psychology, 1300 University Blvd., Birmingham, AL 35294, United States

ARTICLE INFO

Article history:
Received 4 November 2018
Received in revised form 7 April 2019
Accepted 7 April 2019
Available online 22 April 2019

Keywords:
Dynamical systems
Novice driver safety
Driver behavior theory

ABSTRACT

Dynamical systems (DS) theoretical and methodological approaches have furthered our understanding of human development and behavior across many different domains, but have not yet been applied to driver behavior. Using a DS lens, a new theory of driver behavior is proposed, the phase transition framework (PTF), and then applied to understanding how novice drivers develop. The PTF can facilitate new lines of research by providing an individual-level explanatory account that yields the widely observed and poorly understood aggregate declines in novice drivers' crash rates in the early months of licensure. For the purpose of the PTF, learning to drive is defined as changes in the psychological system integral to the acquisition, refinement, automation, and maintenance of competencies necessary for safe driving. In the PTF, the greatest reductions in population-level collision rates observed in the early months of the post-licensure period are hypothesized to be due to qualitative cognitive reorganizations (e.g., adoption of strategies that reduce risk) that cause abrupt shifts in crash risk trajectories into lower risk strata followed by refinement of strategy use and deployment through real-world experience. Critically, the PTF differs from other theories of driver behavior in that it holds that drivers' behavioral development is self-organizing and emergent.

1. Introduction

Motor vehicle crashes persist as a leading cause of death and injury for novice drivers (NCIP). Analyses of population-level crash data indicate that younger age at licensure is associated with higher crash rates, but rates decrease for all age groups fairly rapidly following licensure with the effects of experience outweighing the effects of age on crash involvement (Curry, Pfeiffer, Durbin, & Elliott, 2015; McCartt, Mayhew, Braitman, Ferguson, & Simpson, 2009). Visual depictions of population-level crash rates across license phases in jurisdictions with Graduated Driver Licensing (GDL) programs (i.e., from the supervised learner period to the restricted or intermediate period) illustrate what many researchers, policymakers, and practitioners have come to refer to as the "*learning curve for driving*" (Simons-Morton & Ehsani, 2016; The Institute of Medicine and National Research Council, 2011). The population-level crash rate curves do, in fact, peak at licensure and then bear a striking resemblance to a conventional power law-like learning curve (Foss, Martell, Goodwin, & O'Brien, 2011). This has led many researchers over the past several decades to appeal to power law principles of deliberate practice, to consider the limitations of pre-license driver training and skill transfer, and to conceptually apply the broad literature on the development of expertise towards understanding how novices learn to drive (Foss et al., 2011; Groeger, 2002; Simons-Morton & Ehsani, 2016).

Despite these efforts, the specific processes that lead to reductions in population-level crash rates are still one of the most poorly understood aspects of novice driver safety. In addition, the focus on an incremental-accrual model (i.e., a power law

E-mail address: jhmirman@uab.edu

https://doi.org/10.1016/j.trf.2019.04.010
1369-8478/© 2019 The Author. Published by Elsevier Ltd.

model) of crash risk reduction has narrowly constrained intervention development and driver training programs targeting novice drivers, as well as the evaluations of such interventions and programs. Importantly, the incremental "learning curve" pattern in the population-level crash data can be generated from discontinuous changes (i.e., phase changes) in crash risk at the individual-level (Mirman, Curry, & Mirman, 2019). Researchers in the fields of cognitive and developmental sciences have previously demonstrated that within- and between-subject aggregation can smooth out critically important developmental discontinuities, thereby providing a very misleading picture about underlying change processes (Gray & Lindstedt, 2017; Haider & Frensch, 2002; Seigler, 2006; Thelen & Ulrich, 1991). Therefore, it is important to broaden explanatory models of learning to drive to include the sorts of psycho-behavioral change processes that can sensibly account for the changes in the population-level rates after licensure, and to consider other dimensions and processes of change in addition to those associated with individual differences in learning rates (i.e., slopes).

In the current paper, dynamical systems principles that apply to human ontogeny are reviewed and then leveraged towards a new theory of driver behavior, the phase transition framework (PTF), which is then applied to the special case of the *novice* driver. Crucially, the PTF is a model of change processes, not a model of an end state. Population-level crash rates change a lot during the transition from learner to licensed driver, thus making the novice driver an ideal driver type for this first application of the PTF. It is important to note that a DS perspective has been extensively applied to perceptual-motor control in general (see Warren, 2006), and at least occasionally used to understand the braking behavior of experienced drivers (Warren, 2006; Yilmaz & Warren, 1995). It has not been applied to understanding how novice drivers change, or to driver behavior in general. The evidence on patterns and processes of performance changes among novice drivers is briefly reviewed through a DS lens, and then methodological suggestions to further develop and test the PTF claims are provided.

1.1. Dynamical systems principles

In writing this brief overview, substantial inspiration was taken from other thorough and thoughtful reviews on DS frameworks. Interested readers are referred to these writings for an in-depth overview of DS concepts and their application to cognitive-behavioral aspects of human ontogeny (Granic & Hollenstein, 2003; Lewis, 2000; Smith & Thelen, 2003; Thelen, 2002), physical health and epidemiology (Rickles, Hawe, & Shiell, 2007), as well different DS methodological approaches, and philosophical perspectives (van Geert & Steenbeek, 2005; van Geert, 2003). In addition, a detailed terminological glossary is outside the scope of this paper, and mathematical implications were intentionally not reviewed in order to focus on the conceptual implications of a DS approach. Interested readers are directed to Rickles et al. (2007).

The most important DS principle is that complex systems are softly assembled and self-organizing. That is, patterns of seemingly organized behavior are in fact emergent and generated due to interactions of system sub-components within the context of immediate task demands - not from prescriptions or from the development of schemas. The concept of emergence can be illustrated by using the example case of how young children learn to walk. Smith & Thelen (2003) observed that learning to walk can best be described through a series of phase changes within a dynamical systems framework (Gray & Lindstedt, 2017; Smith & Thelen, 2003). Children do not gradually learn to walk and are not pre-programmed to do so in a specific way. Practically, children need to get around and they try out a few different ways of doing it, typically settle on some version of crawling, stick with it for a few months (i.e., stabilize), some will cruise; then they start walking. During transitions between preferred modalities they might be crawlers and walkers, and they might even be better at crawling than walking (an example of two-phase multi-stability) (Clark & Phillips, 1993), but rather quickly they become pretty good at walking and never seriously use crawling again. As their strength and balance improve, they will, eventually, identify bipedalism as a superior solution to help them get where they want to go (Smith & Thelen, 2003). Children do not learn to walk in a uniform and scripted way, it is highly heterogeneous and non-linear, with the "crawling solution" being emergently generated by some children, but not others (Clark & Phillips, 1993; Smith & Thelen, 2003; Thelen & Ulrich, 1991).

A related and classic developmental example of how environment (or task constraints) influence trajectories of behavioral expression is the "disappearance" of the stepping reflex in infancy. When newborns are held upright in a standing position they will take steps even though they cannot support themselves yet. This reflex usually disappears around 2 months of age, and then is replaced with more mature stepping movement patterns, though babies will still kick in the supine position at the same time as the stepping reflex disappears. Initially, a variety of neurodevelopmental hypotheses were put forward to explain the disappearance of the innate stepping reflex. However, Thelen and Fisher (1982) illustrated that babies can continue to demonstrate the stepping reflex when placed upright in water (Thelen & Fisher, 1982). The stepping reflex "disappeared" not because of scripted neurodevelopmental maturation, but because the infants gained fat tissue more rapidly than muscle tissue making their legs too heavy to lift up for a while, thus the asynchrony between muscle and fat tissue development constrained (i.e., controlled) the stepping reflex. A change in the environment – placing the infants upright in water – allowed the reflex to be observed again. Failure to consider "bottom-up" (e.g., leg weight) and environmental influences (e.g., supine vs. upright positioning; gravity) on behavioral expression misdirected theories of motor development towards overemphasizing "top down" processes (e.g., neurodevelopmental accounts).

Two additional points of clarification highlight the contrast between the DS perspective from prevailing driver behavior theories and from older developmental theories. First, a *complex* system is different from a *complicated* system. For example, theories of driver behavior have tended to be descriptive accounts about the act of driving, which vary in: (a) the number of their constituent constructs and how these constructs are organized, (b) their emphasis on various conscious and unconscious psychological processes (e.g., cognitive, emotional) and biological factors, and (c) the centrality of situational vs. dispositional factors (Fuller, 2005; Kinnear, 2019; Michon, 1985; Ranney, 1994). They are also conceptually deterministic in that the constructs determine driver behavior through various inputs, outputs, and feedback loops. The more constructs and interrelationships, the more *complicated* the theoretical model becomes (i.e., the more parts it has).

In contrast, in a *complex* system it is not the quantity of constituent parts and their connections that matters, but the manner in which the constituent parts interact to produce *emergent* phenomena. Phenomena are described as emergent when they are a property of the whole system, but are not contained in any one constituent part (i.e., a system-level property vs. individual-level property). For example, a double pendulum is not a complicated system - it has only two components (two coupled pendulums) – but it is a complex system because the nonlinear interactions between the two components make it very difficult to predict where the bob (end of the pendulum) will be at any particular time point. A complex system can also be complicated (composed of many parts), but a complicated system need not be complex in the DS sense.

Second, many early theories of human development contained claims that widely observed psychological and behavioral patterns (e.g., robust patterns in motor learning, language acquisition) were generated from pre-programmed scripts, or from uniformly developing schemas (i.e., universal mental models of concepts and their relationships). Claims like these are tempting because developmental patterns can superficially seem purposeful or otherwise directed. They rarely are. Similarly, other theories of development had claims that over-emphasized the importance of environmental influence (e.g., behaviorism). Modern developmental science recognizes that successive organism-environment interactions are the fundamental drivers of both individual variability and commonalities. Put colloquially, it is nature *and* nurture, not nature *or* nurture.

1.2. Studying dynamical systems

To study a dynamical system researchers first need to define its parameters and to identify *collective* variables of interest. These collective variables are a group of variables that result in observable and measurable behavioral patterns, which yield the emergent properties of interest. This is a key concept of emergence; the properties of the whole are not necessarily present in the constituent parts. Other common examples of emergence that readers are likely familiar with include group-level "organized" patterns in bird flocking (e.g., starlings in particular, or geese flying in a "v" shape).

In the physical sciences identifying collective variables is more straightforward because molecules, for example, are much easier to define and to measure than psychological or behavioral constructs. Examples of collective variables studied in the psychological sciences predominantly include motor and cognitive variables like the phasing of alternate stepping to study motor development (Thelen & Ulrich, 1991) and changes in point-of-gaze to study strategy acquisition processes (Stephen, Boncoddo, Magnuson, & Dixon, 2009), but can also include socioemotional development such as changes in the intensity of oppositional behavior (Granic & Hollenstein, 2003). Changes and patterns of the collective variable are studied over multiple time scales in real time (sampling rates spanning milliseconds, minutes) and in developmental time (any organism's unique history). In the context of driving, vehicle kinematics and scanning behavior (e.g., changes in point of gaze) might serve as useful collective variables because they are reflective of more than one subsystem component of driver behavior and are measured using high sampling frequencies essential for detecting fine-grained cognitive-behavioral changes.

Inherent to understanding a dynamical system is evaluating how and why it changes. Dynamical systems researchers use the term state space in reference to contiguous states that the collective variable could possibly occupy. The space of possible states may be very large, but the collective variable will typically only occupy a small number of states in the total space of possible states. This attribute of dynamical systems means that there are prolonged periods of continuity in the collective variable's expression (i.e., no apparent change) even though the system itself is constantly in flux. In state-space diagrams, periods of continuity are often depicted as a basin, which is referred to as a stable, or attractor state (Fig. 1). It is a result of

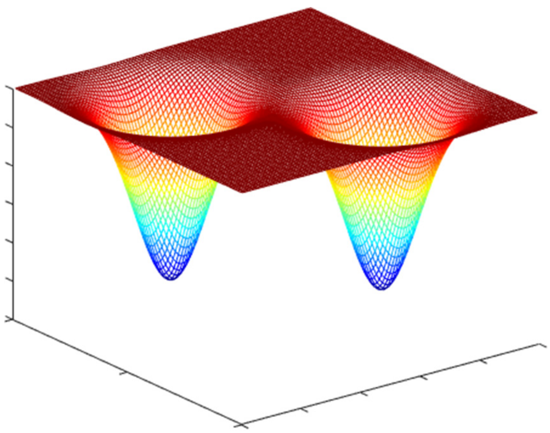

Fig. 1. A system's bifurcation results in two neighboring attractor states.

the ongoing interplay between attractor and repellor forces (e.g., organism or environmental constraints) that behavioral patterns tend to be stable.

Thus, when a collective variable changes states, it does so as a phase shift. These phase shifts can lead to new behavioral patterns through successive bifurcations. Successive bifurcations are often depicted as a series of connected branches splitting into different directions over time, illustrating that the same developmental starting point can lead down many different pathways (i.e., multifinality), and also make it more difficult to get over into another far away "branch". As such, bifurcations open new doors, while they simultaneously shut others. For example, in Fig. 1 we can see a snapshot (in time) of a bifurcation, which is depicted as two basins (i.e., two attractor states). In a hypothetical example we could imagine that a new driver's mood, personality (e.g., neuroticism, agreeableness etc.), and overall emotional intensity (e.g., mild, intense) influences their initial experiences interacting with other drivers (whether they have a lot or a few interactions, the nature of those interactions, etc.); over time these experiences and their feedback may route (or channel) the driver into what we think of as more stable "driving styles". The driving style, then, becomes its own attractor state, limiting the likelihood an aggressive driver will become a conscientious driver. In complex systems such as these, small differences in initial conditions can have large and unexpected downstream consequences, and there are both top-down and bottom-up controlling forces that shape developmental trajectories.

2. A dynamical systems perspective on novice driver behavior: phase transition framework (PTF)

Taking a DS approach to novice driver behavior, and learning to drive specifically, requires giving up the perspective that novice drivers improve incrementally and uniformly as post-license experience accumulates. For the purpose of this paper, "learning to drive" is defined as changes in the psychological system integral to the acquisition, refinement, and automation of action-motor capacities and self-regulatory competencies necessary for safe, goal-directed driving. In this framework, the greatest reductions in population-level collision rates observed in the early months of the post-licensure period are hypothesized to be due to individual-level qualitative cognitive reorganizations (e.g., implicit or explicit adoption of strategies that reduce task demands) that cause relatively abrupt shifts in crash risk trajectories into lower risk strata, followed by refinement of strategy use and deployment through real-world experience. Thus, the topography of the new drivers' state space radically reorganizes from a higher risk state to a lower risk state, but there are comparatively long periods of continuity in the level of risk between these phase changes.

Based on a DS framework, which maintains that mature driver behavioral development is an emergent phenomenon, the PTF of learning to drive has three initial key concepts: (1) changes in crash risk are abrupt, and occur at different times for different novice drivers; (2) person-environment interactions can cause an immediate phase transition, or increase the probability of a future phase transition; and (3) there is substantial heterogeneity with respect to why any given novice driver might be at higher risk of crashing than other novice drivers. Stemming from these concepts, the greatest reductions in crash risk following licensure are due to the implicit or explicit acquisition of strategies that qualitatively reduce the task demands of driving (i.e., strategies that make driving easier and safer) to optimize goal-directed behavior (i.e., the driver gets where he or she wants to go).

2.1. Meaningful changes in crash risk are abrupt, caused by person-environment interactions

Under the DS framework we would expect to see long periods of stability punctuated by discontinuity in drivers' behavior. According to the PTF a driver's development is instigated by the novice wanting to purposefully move (much like the child who wants to walk in the prior example), with development being controlled by specific moment-by-moment changes in task demands that the driver experiences, which create time- and context-dependent opportunities and limits on behavioral expression. Behaviors or capabilities (e.g., speed management) might emerge in a different order depending on the nature of the drivers' environments (e.g., social; physical; vehicular) and their interactions with individual difference factors (e.g., mood; personality). This means that observing behaviors or capabilities is opportunity-dependent.

2.2. Novice drivers' change patterns

Applied to learning to drive, the PTF holds that meaningful reductions in crash risk are due to phase changes, which can be caused by the acquisition of new strategies and insights (implicit or explicit) about the driving task generated to meet changing environmental requirements. Thus far, these claims are directly supported by a computational modeling analysis of population-level crash data and behavioral interventions (Mirman et al., 2019). The analysis demonstrated that both the incremental pattern in the aggregate crash data and the effects of two pre-licensure interventions can be generated from a phase transition framework, that phase transitions can be precipitated by acquiring new strategies such as improved hazard anticipation and identification, and that interventions can cause an immediate phase transition and increase the probability of a future phase transition. Central to the phase transition framework is the idea that the driver behavioral system is continually in flux with prolonged periods of apparent continuity punctuated by comparatively rapid phase-like changes.

Although not analyzed using DS methods, the majority of the novice driver research using in-vehicle data recorders (IVDR)s does show long periods of behavioral continuity among novice driver subgroups that appear to be fairly resistant to change. Analyses of vehicle kinematic data in an observational study (n = 42) over the first 18 months of licensure suggest that new drivers cluster into discernable risk groups: a stable low risk group, a stable high risk group, and a declining risk group (Simons-Morton, Cheon, Guo, & Albert, 2013). An ABA intervention study that used IVDRs to provide feedback to novice drivers (n = 18) within the first year of licensure found that the overall trajectory of kinematic driver errors declined incrementally over the 12-month study period, with some reversal (increase in errors) when the feedback intervention was removed (Carney, McGehee, Lee, Reyes, & Raby, 2010; McGehee, Raby, Carney, Lee, & Reyes, 2007). However, this overall pattern was entirely due to a high error group, which responded to the intervention (reduced kinematic errors); a low risk group entered the study already at low risk and stayed that way (i.e., did not incrementally improve) – they made very few kinematic errors throughout the study, with virtually no change when the feedback intervention was introduced and later removed. A randomized trial of the same IVDR program (n = 90) in which teen drivers were assigned to an immediate feedback condition (a light blinked when they made an error) or an immediate feedback + parental notification condition found that the latter led to significant reductions in kinematic risky driving compared to the immediate feedback only group (Simons-Morton, Bingham, & Ouimet, 2013). Interestingly, the immediate feedback group had a stable pattern of kinematic risky driving during the 13-week study period suggesting that drivers were insensitive to immediate feedback about their driving behavior provided by the device (Simons-Morton et al., 2013). Taken all together, this body of research suggests that novices are *capable* of reducing their risky driving as indexed by vehicle kinematics, but gravitate back to an attractor state (whether it be riskier or safer) in the absence of a strong perturbation (e.g., parental intervention).

Of note, an incremental model of learning predicts that driving performance should improve incrementally for all drivers as they accumulate experience. The "power law of practice" principle further holds that these improvements should proceed more quickly at first and become smaller as overall performance improves. Indeed, these are the core principles of the incremental model of learning to drive safely (Foss et al., 2011; McCartt et al., 2009; Simons-Morton & Ehsani, 2016). Evidence of periods of stability (or plateaus) in novice drivers' learning and performance would directly contradict this hypothesis, because it would constitute evidence of absence of incremental learning and change. Evidence of a different kind of learning and change trajectory, such as abrupt changes in safe driving or long periods of no meaningful change, challenge the incremental model (see Gershon, Ehsani, & Zhu, 2018 for one example). It would be helpful for future research to be conducted on this topic to help identify which facets of learning to drive might follow more conventional power law of practice type patterns (e.g., vehicle handling) vs. those that might be more discontinuous (e.g., hazard anticipation).

2.3. Heterogeneity in risk profiles

The third principle of the PTF stems directly from the second - there is a great deal of heterogeneity in why any given novice could be at an elevated risk of crashing relative to his or her internal baseline, as well as relative to his or her peers. In other words, a novice driver's crash risk is informed by many different channels (Shope & Bingham, 2008), which flux in their influence across time scales (i.e., months to moments) (Lee, 2007). But this does not mean that all of these factors are in-play for any given novice. Most interventions directed towards novice drivers are focused on only reducing one risk factor at a time, which means that most interventions are poorly tuned to the individual needs of the novice drivers comprising the target population. Interventions might be very effective, but only for a small proportion of novices in any given study. Interventions may need to be tuned to each novice's particular risk portfolio, which will change over time.

In Fig. 2 purely hypothetical data are plotted to illustrate this point. Four driver attributes are plotted: (1) basic vehicle control, (2) scanning, (3) sensation seeking, and (4) ability to accurately perceive task demands in different road conditions. In each of the panels the x-axis refers to amount of competence/riskiness, higher values indicate greater competence/less risk. The y-axis refers to the quantity of drivers with this attribute, higher values indicate more drivers (readers can think of them as a newly licensed cohort). In each panel the imagined population distributions are plotted at three time periods relative to licensure: 0 months (red, dotted line), 6 months (green, dashed line), and 12 months (blue, solid line). The vehicle control panel (top left) illustrates that most drivers have strong vehicle control skills at licensure (month 0): basic vehicle control skills are learned quickly and, although there is some variability in the population, they (in this hypothetical scenario) show only minimal improvement after licensure. Scanning (top right) has a different distribution that is evenly distributed at month 0, which gets a lot better on the population-level by 6 months, and little bit better by 12 months. Sensation-seeking (bottom right) is a commonly studied personality variable in the context of risk behavior. In the panel, we can see a skewed distribution with a long tail that evidences very little change over the course of one year. The ability to accurately detect changes in task demands in different road conditions is right skewed at licensure, indicating most novices have low levels of skill in this domain. Six-months later the distribution is relatively normal and has shifted, evidencing population-level improvement. At 12 months this same pattern has continued such that the distribution is now left skewed. All four attributes vary across the population and at least three show some degree of improvement (i.e., reduction in crash risk) after licensure, but some change more than others, and at different trajectories. Not depicted is how these four attributes might co-vary with one another at the individual-level.

Consider three hypothetical teen drivers who drive roughly the same amount in the same conditions for the first 12 months of licensure. Teen Driver A has strong vehicle control skills at months 0, 6, and 12 (95th percentiles), has moderate scanning skills at months 0 and 6 (50th percentiles), and then improves to the 85% at month 12, is high on sensation seeking (95th percentiles at months 0, 6, and 12), and also has strong task demand detection skills (90th percentiles at

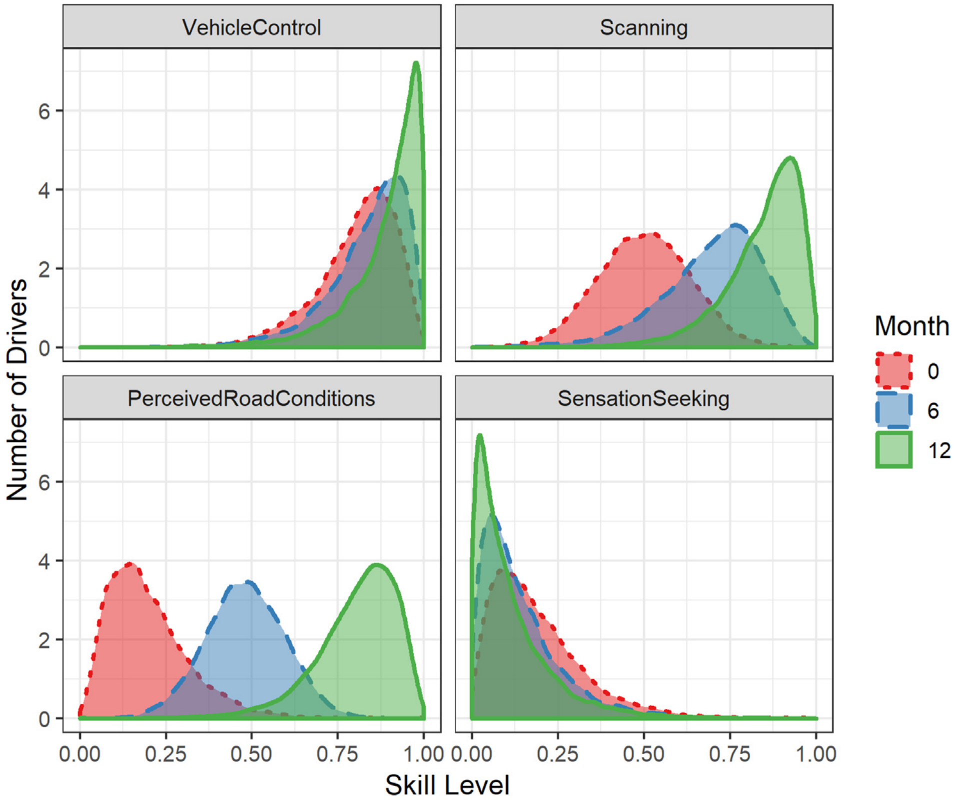

Fig. 2. Hypothetical patterns of heterogeneity of driver attributes across time.

months 0, 6, and 12). Driver A may benefit most from strong parental limit-setting and Graduated Driver Licensing provisions to mitigate opportunities for his or her sensation seeking to lead to crash involvement, and hazard perception training.

Driver B has poor vehicle maneuvering skills that moderately improve over time 0, 6, and 12 months (10th percentile, 20th percentile, and 30th percentile respectively), a similar pattern of scanning ability (12th percentile, 18th percentile, and 30th percentile respectively), is a low sensation seeker (5th percentile across all three periods), but has strong task demand detection ability across time 85th, 90th, and 90th from 0 to 12 months. Driver B might benefit the most from training in vehicle handling. While most drivers do not have this problem, it is a problem for Driver B, and should Driver B ever get in to a crash, it may be most likely due to this driver's deficit in maneuvering.

Finally, Driver C has average vehicle control, scanning, and task demand ability (50th percentile) and is a low sensation seeker (20th percentile) at licensure. However, Driver C does not improve with experience; 12 months later he or she has dropped to the 10th percentile with respect to vehicle control, scanning, and task demand skills, but retained his or her sensation seeking tendency relative to his or her peers (20th percentile). Driver C might need targeted interventions in each of the domains as well as an assessment to understand why he or she is failing to make progress (i.e., unable to learn from experience).

As it stands, we lack a sufficiently detailed and mechanistic understanding of how different driving abilities are acquired, what typically vs. atypically developing drivers look like, which interventions would be most successful for specific drivers at which time points, and which factors are most malleable. This lack of information has made it incredibly difficult to develop effective individual-level behavioral interventions. Moreover, the assumption that every driver in the study population needs the intervention may make it close to impossible to detect a meaningful effect. Researchers should assume effect heterogeneity for any behavioral intervention and plan their research study and statistical plans accordingly *prior* to starting the study. Having a strong theory about why the intervention should work and why the effect might vary by individual difference factors is crucial for guiding both the study design and the analytical plan to ensure that the study is appropriately powered.

3. Methodological considerations for testing PTF claims

There is a very limited knowledge base on recently licensed novices' driving performance trajectories due to the lack of studies on this topic. Between-group comparisons are the most common (e.g., experienced drivers vs. younger, novice drivers of different age ranges and amount of inexperience). Crash databases are often used to examine changes in population-level crash patterns and types of collisions (Braitman, Kirley, McCartt, & Chaudhary, 2008; Foss et al., 2011), but these data cannot be used to make strong inferences about individual-level change processes. Although some self-report longitudinal survey studies exist, they are not appropriate for capturing DS properties due to the very low sampling rate and inability to capture collective variables of interest (Roman, Poulter, Barker, McKenna, & Rowe, 2015).

Detecting phase changes requires a well-defined collective variable with: (a) fine-grained measurements, such as motion tracking or eye tracking with high sampling frequencies, (b) clear, testable ideas about what phenomena will be observed when the participant has transitioned from one phase to another (e.g., if they start using a new strategy), and (c) clear, testable ideas about the kinds of experiences that can initiate a transition. It is up to individual researchers to specify what these might be. In addition, detecting phase transitions may also require new or different measures or analytic strategies. Discontinuous patterns of change are common in ontogeny, but they can be obscured: (1) when measurements are taken too infrequently (i.e., with wide sampling intervals), (2) by averaging within individuals (e.g., averaging across time; averaging across trials) (Seigler, 2006; Siegler & Shipley, 1995; Siegler, 1987), and (3) by averaging across individuals (Haider & Frensch, 2002; Murre & Chessa, 2011). Finally, researchers must pivot from thinking about variability as "noise" towards viewing variability as evidence of important change processes associated with a complex system self-organizing to meet changing task demands. Each of these issues is elaborated on below.

3.1. Sampling interval selection

Sampling interval selection is critical in any study of change processes. Because dynamical systems are constantly in flux and are non-linear they require study designs with high sampling frequencies. Assessing dynamical systems typically entails hundreds or thousands of samples per individual time series on the level of 2–4 orders of magnitude higher than the phenomena being measured. The critical point is that sampling intervals should be directly informed by the developmental timescale of the dynamical process one is trying to measure.

Because the PTF holds that meaningful developmental changes can occur quickly, methodologies with wide sampling intervals are not well-suited to testing PTF claims. This would immediately exclude most survey-based cohort studies that have months-long increments between assessments and are not assessing a collective variable. Data from studies using unobtrusive measures with high sampling frequencies such as in-vehicle technologies could potentially provide an excellent window into changes in driver performance that could be associated with stability and change in drivers' crash risk, if we interpret kinematic risky driving as a suitable collective variable. However, very few studies have been conducted (see (Ehsani et al., 2017), and none of these data have been analyzed using a DS perspective (Ehsani et al., 2017; Simons-Morton, Ehsani, Gershon, Klauer, & Dingus, 2017). For example, the aggregation of finely sampled measures into months-long time bins would be inconsistent with a DS approach.

3.2. Alternatives to central tendency measures

In physical systems transitions from one stable state to another stable state are characterized by a brief intermediate period of instability and increased entropy. Similarly, the spontaneous discovery of new strategies can exhibit characteristic transient changes in entropy of hand or eye movements (Dixon, Stephen, Boncoddo, & Anastas, 2010; Stephen et al., 2009). This means that phase shifts can be detected by assessing the collective variable using a variety of entropy measures that are well-suited for psycho-physiological variables, such as determining if peak and prior entropy levels change across performance trials on an experimental task (Riley & Holden, 2012; Stephen et al., 2009). In the novice driver research field, most researchers implicitly prioritize central tendency measures without strong consideration of other aspects of an individual participant's distribution or groups of participants' distributions on variables of interest. Yet, in the context of wanting to understand phase transitions, an increase in variability may be an indicator of crucial information about change processes and conventional approaches to smoothing out data can make this information disappear. Aggregation is discussed in more detail in the subsequent section.

Behaviorally, researchers can assess for dips followed by leaps in performance owing to idea that behavioral performance may worsen slightly while learners are exploring new strategies (Gray & Lindstedt, 2017). Analytic and data visualization approaches (e.g., recurrence quantification analysis; state space grids) can be found in the dynamical systems literature (Erickson, Côté, Hollenstein, & Deakin, 2011; Granic & Hollenstein, 2003) and in classic developmental studies that utilize small n study designs with high sampling rates (Seigler, 2006). Small n study designs may be uniquely helpful for detecting phase changes due to the emphasis on individual time series data (Smith & Little, 2018). These designs offer an added practical benefit in that they can retain strong statistical power but without the expense and complexity associated with larger sample sizes, such as those associated with population-based surveys or larger scale observational studies (Smith & Little, 2018). When conducted in conjunction with computational cognitive modeling methods (Sun, 2009), which provide a methodologically rigorous approach for instantiating theoretical claims, small n designs can be especially useful.

3.3. Aggregation considerations

Averaging *a participant's* data over time (or trials), as well as averaging over *multiple participants'* data can produce patterns that do not reflect underlying psychological processes. While this is broadly true, it is a particular problem in the case of a PTF that stipulates that changes will be relatively abrupt, and consequently, that there will be groups of drivers in different phases. For example, prior research on strategy use has shown that children and adults often use multiple strategies when solving problems and may use different strategies over different experimental trials in the same study. Averaging participants' performance across trials when they are using different strategies on a per-trial basis can provide misleading information about their strategy use and success of each type of strategy in terms of its absolute and relative speed and accuracy (Siegler, 1987). Moreover, as stated previously, learners can try out several strategies before settling on an optimal one, and during this time of strategy exploration they might perform worse on a task. Averaging in this context can mask the leaps or gains in performance relative to a gradual trend line, as well as any preceding dips or drops in performance (Gray & Lindstedt, 2017).

Averaging across individuals can also be problematic because the timing of phase transitions can differ across individuals. For example, a substantial amount of attention has been focused on ascertaining if quantities of pre-license practice driving are associated with post-license crash risk, which is a very reasonable question to ask. There is not strong and consistent evidence that the quantity of pre-license experience (number supervised practice driving hours) reduces post-licensure MVC involvement (Mirman, Curry, Elliott, Long, & Pfeiffer, 2018; Simons-Morton & Ehsani, 2016; Williams, 2017). The absence of strong associations between the number of supervised practice hours during the learner period and crash risk holds across cultures and licensing programs irrespective of study design (Williams, 2017). It may be possible to explain the lack of strong evidence of an association between practice quantity and crash risk due to the components of practice quality (e.g., diversity, challenge, instructional quality; see Mirman & Kay, 2012) not being routinely measured and, relatedly, by appealing to the principle of transfer-appropriate processing.

Basic and applied research in cognitive psychology has shown that generalization of learning depends on similarity between the learning context(s) and the testing context(s) (Adams, 1987; Groeger & Banks, 2007). If supervised practice driving is insufficiently similar to independent driving, then it is plausible that any learning that happens during those practice drives would not transfer to independent driving. Whether elaborated with transfer-appropriate processing or not (Groeger & Banks, 2007), most novice drivers are not involved in a police-reported crash during the first year of independent driving. This could suggest that there is at least some proportion of drivers who are exiting the learner period at lower risk than their counterparts – although we cannot infer that the absence of a crash is necessarily indicative of having learned to drive safely. There are many reasons why a driver may not have crashed other than increased competence.

One PTF interpretation of the lack of association between practice quantity and post-license outcomes is that learning during the "learner period" is discontinuous, and the *quantity* of practice hours is a poor proxy for whether the discontinuity has occurred or not. Summing, or otherwise aggregating, practice hours, therefore, would blur the timing of that discontinuous change in performance, making it harder to observe an effect of pre-license practice on post-license outcomes. In other words, if what matters most about practice is if it induces phase transitions or not, which can occur at different times for different learners, benefits of practice would be extremely difficult to observe if within and between-group aggregation-type analyses are used. When we aggregate practice quantity, however it is captured (e.g., miles driven, hours practiced), we lose the ability to examine processes of change and are limited to searching for a *common or average* threshold of practice that should be universally beneficial for all learners. As most transportation scientists agree, there is no theoretical motivation for why a common "magical" threshold of practice should exist.

4. Summary and future directions

In this paper, a DS lens was applied to understanding driver behavior and the special case of the novice driver was used to explore how a DS theoretical and methodological approach could help to address important unanswered questions about how to safely expedite novices' acquisition of driving competence. Although there are many implications for research and practice, three are paramount.

First, drivers' behavior is emergent in that it is a product of the interactions among sub-system components – the whole will always be greater than the sum of its parts. Thus, traditional theoretical and methodological approaches of decomposing drivers' behavior and isolating subsystem components is inherently problematic. Once the behaviors are isolated they are removed from the interactive properties that yield the behavior we are interested in. We cannot predict the trajectory of a coupled pendulum by studying the trajectories of two single pendulums. Similarly, we cannot predict how novices learn to drive by studying any one sub component of novice drivers' cognition and behavior. This is why defining collective variables is so important.

Second, the PTF holds that there will be different groups of novices at varying levels of risk throughout the licensure period, and as reviewed in this paper, the available evidence suggests that this is indeed the case. Failure to take subgroup-level risk into account when designing or evaluating driver behavioral studies of any scale can have significant consequences on the conclusions that are drawn, especially in the context of correlational or experimental designs that involve computing an average effect size. To some extent, this issue (that of statistical confounding by unmeasured subgroups), can be thought of as conceptual extension of Simpson's Paradox, which is a data pooling and inference problem that can occur when data from subgroups are combined (Fig. 3) (Kievit, Frankenhuis, Waldorp, & Borsboom, 2013).

In the aggregate, misleading patterns among variables can appear, which then disappear or reverse when the data are disaggregated into *meaningful* subgroups. For example, in Fig. 3 the overall correlation between x and y is a positive correlation of 0.42, but the correlations between x and y within the subgroups are negative -0.41 and -0.43. In a hypothetical

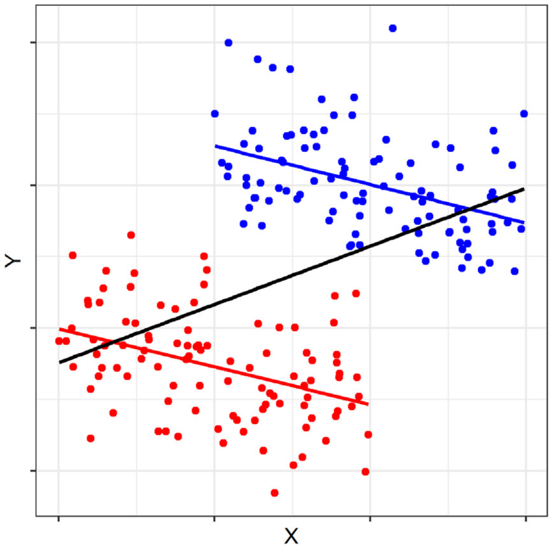

Fig. 3. Illustration of Simpson's Paradox.

example, we could imagine that novices' driving violations are plotted on the y-axis and time-licensed is plotted on the x-axis; drivers who have not had a phase transition (in any domain) are represented by the blue dots (higher cluster) and drivers who have are represented by the red dots (lower cluster). The subgroups could also represent common individual difference variables such as higher (blue) and lower (red)[1] mileage drivers or gender (male, female), for just a few examples. The problem occurs when one generalizes the pattern from the aggregate (population) to individuals or to subgroups. In the hypothetical data in Fig. 3, it is statistically accurate to say that drivers' violations increase with time-licensed at the population-level as observed in the novice driver literature (McKenna, 2018), but the opposite pattern between violations and time licensed might exist at the group level. The issue of confounding by unmeasured subgroups might account for some of the more conflicting, paradoxical, and counter-intuitive findings concerning novice drivers to date (Kinnear, 2019).

Third, on the topic of inferences, researchers know that observing a correlation between two variables in a population or a convenience sample of participants does not mean that there is a causal relationship between those two variables at the individual-level (e.g., Fig. 3). However, the practice of using surveys administered to some sample of novice drivers and then using that pattern of correlations to develop individual-level behavioral interventions relies on the assumption of causality. Rarely is the individual-level causality assumption tested by researchers prior to intervention development and evaluation. The PTF specifically, and a DS approach more generally, provides a family of new approaches to evaluate individual-level change processes prior to the development and evaluation of individual-level interventions with a small number of participants. This new family of methods could help to prevent wasted resources on developing and evaluating interventions that are intended to change causal relationships that do not exist.

In conclusion, while the initial focus of the PTF was on novice drivers it can just as easily be applied to questions about experienced driver behavior, such as understanding behavioral adaptation, and everyday driving. It is likely that there is quite a bit of existing data available that could be re-analyzed using the methodological approaches outlined in the current manuscript. Additional empirical evidence is critical for evaluating the PTF's claims and revising the theory in light of new findings.

Contribution statements

J. Hafetz Mirman conceived of and drafted the manuscript, critically revised the manuscript, and approved its submission for peer review.

Funding & acknowledgements

Funding for this work, specifically for the time and effort of the author, was provided by The University of Alabama at Birmingham. The author is grateful for thoughtful feedback on the manuscript from Emma Goodman and Dan Mirman, for assistance with data visualization from Dan Mirman, and for insightful comments about driver behavior theory from Neale Kinnear.

[1] For interpretation of color in Figs. 3 the reader is referred to the web version of this article.

Conflicts of interest

The author has no conflicts of interest to disclose in relation to the manuscript.

References

Adams, J. A. (1987). Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychological Bulletin, 101*(1), 41–74. https://doi.org/10.1037/0033-2909.101.1.41.

Braitman, K. A., Kirley, B. B., McCartt, A. T., & Chaudhary, N. K. (2008). Crashes of novice teenage drivers: Characteristics and contributing factors. *Journal of Safety Research, 39*(1), 47–54. https://doi.org/10.1016/j.jsr.2007.12.002.

Carney, C., McGehee, D. V., Lee, J. D., Reyes, M. L., & Raby, M. (2010). Using an event-triggered video intervention system to expand the supervised learning of newly licensed adolescent drivers. *American Journal of Public Health, 100*(6), 1101–1106. https://doi.org/10.2105/AJPH.2009.165829.

Clark, J. E., & Phillips, S. J. (1993). A longitudinal study of intralimb coordination in the first year of independent walking: A dynamical systems analysis. *Child Development, 64*(4), 1143–1157.

Curry, A. E., Pfeiffer, M. R., Durbin, D. R., & Elliott, M. R. (2015). Young driver crash rates by licensing age, driving experience, and license phase. *Accident Analysis & Prevention, 80*, 243–250. https://doi.org/10.1016/j.aap.2015.04.019.

Dixon, J. A., Stephen, D. G., Boncoddo, R., & Anastas, J. (2010). The self-organization of cognitive structure. *Psychology of Learning and Motivation, 52*, 343–384.

Ehsani, J. P., Klauer, S., Zhu, C., Gershon, P., Dingus, T. A., & Simons-Morton, B. G. (2017). Naturalistic assessment of the learner license period. *Accident Analysis & Prevention, 106*(June), 275–284. https://doi.org/10.1016/j.aap.2017.06.014.

Erickson, K., Côté, J., Hollenstein, T., & Deakin, J. (2011). Examining coach-athlete interactions using state space grids: An observational analysis in competitive youth sport. *Sport, Exercise, and Performance Psychology, 12*(6), 645–654. https://doi.org/10.1016/j.psychsport.2011.06.006.

Foss, R. D., Martell, C. A., Goodwin, A. H., & O'Brien, N. P. (2011). *Measuring changes in teenage driver crash characteristics during the early months of driving*. Washington, DC: AAA Foundation for Traffic Safety.

Fuller, R. (2005). Towards a general theory of driver behaviour. *Accident Analysis & Prevention, 37*(3), 461–472. https://doi.org/10.1016/j.aap.2004.11.003.

Gershon, P., Ehsani, J. P., Zhu, C., et al (2018). Crash risk and risky driving behavior among adolescents during learner and independent driving periods. *Journal of Adolescent Health, 63*(5), 568–574. https://doi.org/10.1016/j.jadohealth.2018.04.012.

Granic, I., & Hollenstein, T. (2003). Dynamic systems methods for models of developmental psychopathology. *Development and Psychopathology, 15*(3), 641–669. https://doi.org/10.1017/S0954579403000324.

Gray, W. D., & Lindstedt, J. K. (2017). Plateaus, dips, and leaps: Where to look for inventions and discoveries during skilled performance. *Cognitive Science, 41*(7), 1838–1870. https://doi.org/10.1111/cogs.12412.

Groeger, J. A. (2002). Trafficking in cognition: Applying cognitive psychology to driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 5*(4), 235–248. https://doi.org/10.1016/S1369-8478(03)00006-8.

Groeger, J. A., & Banks, A. P. (2007). Anticipating the content and circumstances of skill transfer – Unrealistic expectations of driver training and graduated licensing? *Ergonomics, 50*(8), 1250–1263.

Haider, H., & Frensch, P. A. (2002). Why aggregated learning follows the power law of practice when individual learning does not: comment on Rickard (1997, 1999), Delaney et al. (1998), and Palmeri (1999). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(2), 392–406. https://doi.org/10.1037/0278-7393.28.2.392.

Kievit, R. A., Frankenhuis, W. E., Waldorp, L. J., & Borsboom, D. (2013). Simpson's paradox in psychological science: A practical guide. *Frontiers in Psychology, 4*(AUG), 1–14. https://doi.org/10.3389/fpsyg.2013.00513.

Kinnear N. (2019) Evolution of driver behavior theories. In Driver behavior in an era of new technologies: Implications for theory and training. Washington DC, United States: Transportation Research Board Annual Meeting.

Kinnear N. (2019). Challenging our assumptions: a consideration of selected recent evidence from young and novice driver research. In The struggle is real: Interventions for novice teen drivers. Washington, D.C.: Transportation Research Board Annual Meeting.

Lee, J. D. (2007). Technology and teen drivers. *Journal of Safety Research, 38*(2), 203–213. https://doi.org/10.1016/j.jsr.2007.02.008.

Lewis, M. D. (2000). The promise of dynamic systems approaches for an integrated account of human development. *Child Development, 71*(1), 36–43. https://doi.org/10.1111/1467-8624.00116.

McCartt, A. T., Mayhew, D. R., Braitman, K. A., Ferguson, S. A., & Simpson, H. M. (2009). Effects of age and experience on young driver crashes: Review of recent literature. *Traffic Injury Prevention, 10*(3), 209–219. https://doi.org/10.1080/15389580802677807.

McGehee, D. V., Raby, M., Carney, C., Lee, J. D., & Reyes, M. L. (2007). Extending parental mentoring using an event-triggered video intervention in rural teen drivers. *Journal of Safety Research, 38*(2), 215–227. https://doi.org/10.1016/j.jsr.2007.02.009.

McKenna, F. P. (2018). Do low levels of confidence suppress the tendency to violate? *Transportation Research Part F: Traffic Psychology and Behaviour, 55*, 136–140. https://doi.org/10.1016/j.trf.2018.03.004.

Michon, J. A. A. (1985). In *Critical view of driver behavior models: What do we know, what should we do?* https://doi.org/10.1007/978-1-4613-2173-6.

Mirman, J. H., Curry, A. E., Elliott, M. R., Long, L., & Pfeiffer, M. R. (2018). Can adolescent drivers' motor vehicle crash risk be reduced by pre-licensure intervention? *Journal of Adolescent Health, 62*(3), 341–348. https://doi.org/10.1016/j.jadohealth.2017.09.015.

Mirman, J. H., Curry, A. E., & Mirman, D. (2019). Learning to drive: A reconceptualization. *Transportation Research Part F: Traffic Psychology and Behaviour, 62*, 316–326. https://doi.org/10.1016/j.trf.2019.01.010.

Mirman, J. H., & Kay, J. (2012). From passengers to drivers: Parent perceptions about how adolescents learn to drive. *Journal of Adolescent Research, 27*(3). https://doi.org/10.1177/0743558411409934.

Murre, J. M. J., & Chessa, A. G. (2011). Power laws from individual differences in learning and forgetting: Mathematical analyses. *Psychonomic Bulletin & Review, 18*(3), 592–597. https://doi.org/10.3758/s13423-011-0076-y.

National Center for Injury Prevention and Control (NCIP). Web-based Injury Statistics Query and Reporting System (WISQARS).

Ranney, T. A. (1994). Models of driving behavior: A review of their evolution. *Accident Analysis & Prevention, 26*(6), 733–750. https://doi.org/10.1016/0001-4575(94)90051-5.

Rickles, D., Hawe, P., & Shiell, A. (2007). A simple guide to chaos and complexity. *Journal of Epidemiology and Community Health, 61*(11), 933–937. https://doi.org/10.1136/jech.2006.054254.

Riley, M. A., & Holden, J. G. (2012). Dynamics of cognition. *Wiley Interdisciplinary Reviews: Cognitive Science, 3*(6), 593–606. https://doi.org/10.1002/wcs.1200.

Roman, G. D., Poulter, D., Barker, E., McKenna, F. P., & Rowe, R. (2015). Novice drivers' individual trajectories of driver behavior over the first three years of driving. *Accident Analysis & Prevention, 82*, 61–69. https://doi.org/10.1016/j.aap.2015.05.012.

Seigler, R. S. (2006). Microgenetic analysis of learning. In D. Kuhn, R. S. Seigler, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology* (6th ed., pp. 464–510). Wiley.

Shope, J., & Bingham, C. (2008). Teen driving: motor-vehicle crashes and factors that contribute. *American Journal of Preventive Medicine, 35*(3), S261–S271. https://doi.org/10.1016/j.amepre.2008.06.022.

Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General, 116*(3), 250–264. https://doi.org/10.1037/0096-3445.116.3.250.

Siegler, R. S., & Shipley, C. (1995). In T. Simon & G. Halford (Eds.), *Variation, selection, and cognitive change*. Hillsdale, NJ: Erlbaum.

Simons-Morton, B. G., Bingham, R., Ouimet, M. C., et al (2013). The effect on teenage risky driving of feedback from a safety monitoring system: A randomized controlled trial. *Journal of Adolescent Health, 53*(1), 21–26. https://doi.org/10.1016/j.jadohealth.2012.11.008.

Simons-Morton, B. G., Cheon, K., Guo, F., & Albert, P. (2013). Trajectories of kinematic risky driving among novice teenagers. *Accident Analysis & Prevention, 51*, 27–32. https://doi.org/10.1016/j.aap.2012.10.011.

Simons-Morton, B. G., & Ehsani, J. P. (2016). Learning to drive safely: Reasonable expectations and future directions for the learner period. *Safety, 2*, 1–19. https://doi.org/10.3390/safety2040020.

Simons-Morton, B. G., Ehsani, J. P., Gershon, P., Klauer, S., & Dingus, T. (2017). Teen driving risk and prevention: Naturalistic driving research contributions and challenges. *Safety, 3*(4), 29. https://doi.org/10.3390/safety3040029.

Smith, P. L., & Little, D. R. (2018). Small is beautiful: In defense of the small-N design. *Psychonomic Bulletin & Review, 25*(6), 2083–2101. https://doi.org/10.3758/s13423-018-1451-8.

Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Sciences, 7*(8), 343–348. https://doi.org/10.1016/S1364-6613(03)00156-6.

Stephen, D. G., Boncoddo, R. A., Magnuson, J. S., & Dixon, J. A. (2009). The dynamics of insight: Mathematical discovery as a phase transition. *Memory & Cognition, 37*(8), 1132–1149. https://doi.org/10.3758/MC.37.8.1132.

Sun, R. (2009). Theoretical status of computational cognitive modeling. *Cognitive Systems Research, 10*(2), 124–140. https://doi.org/10.1016/j.cogsys.2008.07.002.

The Science of Adolescent Risk-Taking: Workshop Report. Institute of Medicine (US) and National Research Council (US) Committee on the Science of Adolescence. Washington (DC): National Academies Press (US); 2011. ISBN-13: 978-0-309-15852-7 – ISBN-10: 0-309-15852-4.

Thelen, E., & Fisher, D. M. (1982). Newborn stepping: An explanation for a ''disappearing" reflex. *Developmental Psychology, 15*, 447–453.

Thelen, E., & Ulrich, B. D. (1991). Hidden skills: A dynamic systems analysis of treadmill stepping during the first year. *Monographs of the Society for Research in Child Development, 56*(1), 1–98. https://doi.org/10.2307/1166099.

Thelen E. (2002). Self-organization in developmental processes: Can systems approaches work? In M. Johnson, Y. Munkata, and R. O. Gilmore (Eds.) Brain development and cognition: A reader. (Vol. 22, 2nd ed., pp. 337–374). http://doi.org/10.1002/9780470753507.ch18.

van Geert, P. (2003). Dynamic systems approaches and modeling of developmental processes. *Handbook of Developmental Psychology*, 640–672. https://doi.org/10.4135/9781848608306.

van Geert, P., & Steenbeek, H. (2005). Explaining after by before: Basic aspects of a dynamic systems approach to the study of development. *Developmental Review, 25*(3–4), 408–442. https://doi.org/10.1016/j.dr.2005.10.003.

Warren, W. H. (2006). The dynamics of perception and action. *Psychological Review, 113*(2), 358–389. https://doi.org/10.1037/0033-295X.113.2.358.

Williams, A. F. (2017). Graduated driver licensing (GDL) in the United States in 2016: A literature review and commentary. *Journal of Safety Research, 63*, 29–41. https://doi.org/10.1016/j.jsr.2017.08.010.

Yilmaz, E. H., & Warren, W. H. (1995). Visual control of braking: A test of the r hypothesis. *Journal of Experimental Psychology: Human Perception and Performance, 21*(5), 996–1014.